

Week 2 Exercises

Nelson Tran

July 16, 2023

Please complete all exercises below. You may use stringr, lubridate, or the forcats library.

Place this at the top of your script: library(stringr) library(lubridate) library(forcats)

Exercise 1

Read the sales_pipe.txt file into an R data frame as sales.

```
# Your code here'

sales_pipe <- read.delim("Data/sales_pipe.txt"
                        ,stringsAsFactors=FALSE
                        ,sep = "|"
                        ,fileEncoding = "WINDOWS-1252"
                        )
```

Exercise 2

You can extract a vector of columns names from a data frame using the colnames() function. Notice the first column has some odd characters. Change the column name for the FIRST column in the sales data frame to Row.ID.

Note: You will need to assign the first element of colnames to a single character.

```
# Your code here

colnames(sales_pipe)

## [1] "ï...Row.ID" "Order.ID" "Order.Date" "Ship.Date"
## [5] "Ship.Mode" "Customer.ID" "Customer.Name" "Segment"
## [9] "Country" "City" "State" "Postal.Code"
## [13] "Region" "Product.ID" "Category" "Sub.Category"
## [17] "Product.Name" "Sales" "Quantity" "Discount"
## [21] "Profit"

colnames(sales_pipe)[1] <- "Row.ID"
```

Exercise 3

Convert both Order.ID and Order.Date to date vectors within the sales data frame. What is the number of days between the most recent order and the oldest order? How many years is that? How many weeks?

Note: Use lubridate

```

# Your code here
sales_pipe$Ship.Date = as.Date(sales_pipe$Ship.Date,
                              format = '%B %d %Y'
                              )

sales_pipe$Order.Date <- as.Date(sales_pipe$Order.Date
                                ,format='%m/%d/%Y'
                                )

#I believe you can use the function summary too, that would give you the min and max ONLY after you cha
summary(sales_pipe)

```

```

##      Row.ID      Order.ID      Order.Date      Ship.Date
## Min.   :    1  Length:4928  Min.   :2014-01-03  Min.   :2014-01-07
## 1st Qu.:2322  Class :character 1st Qu.:2015-06-11 1st Qu.:2015-06-15
## Median :4966  Mode  :character Median :2016-07-17 Median :2016-07-22
## Mean   :5031  Mean   :2016-05-09 Mean   :2016-05-13
## 3rd Qu.:7666  3rd Qu.:2017-05-28 3rd Qu.:2017-06-02
## Max.   :9994  Max.   :2017-12-30 Max.   :2018-01-05
## Ship.Mode      Customer.ID      Customer.Name      Segment
## Length:4928      Length:4928      Length:4928      Length:4928
## Class :character  Class :character  Class :character  Class :character
## Mode  :character  Mode  :character  Mode  :character  Mode  :character
##
##
##      Country      City      State      Postal.Code
## Length:4928      Length:4928      Length:4928      Min.   : 1040
## Class :character  Class :character  Class :character  1st Qu.:22204
## Mode  :character  Mode  :character  Mode  :character  Median :60076
##                                     Mean   :55283
##                                     3rd Qu.:90008
##                                     Max.   :99301
##      Region      Product.ID      Category      Sub.Category
## Length:4928      Length:4928      Length:4928      Length:4928
## Class :character  Class :character  Class :character  Class :character
## Mode  :character  Mode  :character  Mode  :character  Mode  :character
##
##
##      Product.Name      Sales      Quantity      Discount
## Length:4928      Min.   :    0.444  Min.   : 1.000  Min.   :0.0000
## Class :character  1st Qu.:   16.962  1st Qu.: 2.000  1st Qu.:0.0000
## Mode  :character  Median :   53.810  Median : 3.000  Median :0.2000
##                                     Mean   : 233.613  Mean   : 3.761  Mean   :0.1592
##                                     3rd Qu.: 211.960  3rd Qu.: 5.000  3rd Qu.:0.2000
##                                     Max.   :22638.480  Max.   :14.000  Max.   :0.8000
##      Profit
## Min.   : -3839.990
## 1st Qu.:    1.678
## Median :    8.495
## Mean   :   25.776
## 3rd Qu.:   29.538
## Max.   :  8399.976

```

```
max(sales_pipe$Order.Date)
```

```
## [1] "2017-12-30"
```

```
min(sales_pipe$Order.Date)
```

```
## [1] "2014-01-03"
```

```
time_1 <- "2017-12-30"
```

```
time_2 <- "2014-01-03"
```

```
difftime(time_1, time_2, units = "days")
```

```
## Time difference of 1457 days
```

```
#difftime can't use years as an arg so I used weeks and divided by 52.25 (accounting for leap year) whi  
difftime(time_1, time_2, units = "weeks")/52.25
```

```
## Time difference of 3.983595 weeks
```

```
difftime(time_1, time_2, units = "weeks")
```

```
## Time difference of 208.1429 weeks
```

Exercise 4

What is the average number of days it takes to ship an order?

```
# Your code here
```

```
mean(sales_pipe$Ship.Date - sales_pipe$Order.Date)
```

```
## Time difference of 3.908482 days
```

Exercise 5

How many customers have the first name Bill? You will need to split the customer name into first and last name segments and then use a regular expression to match the first name bill. Use the `length()` function to determine the number of customers with the first name Bill in the sales data.

```
# Your code here
```

```
unique(sales_pipe$Customer.Name)
```

```
##      [1] "Claire Gute"           "Darrin Van Huff"      "Sean O'Donnell"  
##      [4] "Brosina Hoffman"      "Andrew Allen"         "Irene Maddox"  
##      [7] "Harold Pawlan"        "Pete Kriz"            "Ken Black"  
##     [10] "Sandra Flanagan"      "Emily Burns"          "Eric Hoffmann"  
##     [13] "Tracy Blumstein"      "Ken Brennan"          "Joel Eaton"  
##     [16] "Stewart Carmichael"   "Duane Noonan"         "Julie Creighton"  
##     [19] "Christopher Schild"    "Patrick O'Donnell"     "Paul Gonzalez"  
##     [22] "Gary Mitchum"         "Jim Sink"              "Karl Braun"  
##     [25] "Janet Martin"         "Pete Armstrong"       "Cynthia Voltz"  
##     [28] "Clay Ludtke"          "Sally Hughsby"        "Sandra Glassco"  
##     [31] "Helen Andreada"       "Maureen Gastineau"    "Joni Blumstein"  
##     [34] "David Smith"          "Valerie Dominguez"    "Erin Ashbrook"  
##     [37] "David Bremer"         "Ken Lonsdale"         "Dianna Wilson"
```

## [40]	"Logan Haushalter"	"Kelly Collister"	"Delfina Latchford"
## [43]	"Dan Reichenbach"	"Alan Haines"	"Nick Zandusky"
## [46]	"Kelly Lampkin"	"Jonathan Doherty"	"Jonathan Howell"
## [49]	"Alan Schoenberger"	"Corey Roper"	"Shahid Hopkins"
## [52]	"Ben Peterman"	"Thomas Seio"	"Andy Gerbode"
## [55]	"Sung Pak"	"Nathan Mautz"	"Frank Atkinson"
## [58]	"Grace Kelly"	"Don Jones"	"Patrick O'Brill"
## [61]	"John Lucas"	"Doug Bickford"	"Alyssa Crouse"
## [64]	"Chad Sievert"	"Clay Cheatham"	"Tamara Dahlen"
## [67]	"Adam Bellavance"	"Jeremy Lonsdale"	"Heather Kirkland"
## [70]	"Victoria Brennan"	"Katrina Willman"	"Julia Dunbar"
## [73]	"Michael Kennedy"	"Guy Thornton"	"Arthur Gainer"
## [76]	"Muhammed MacIntyre"	"Allen Rosenblatt"	"Alejandro Savely"
## [79]	"Laura Armstrong"	"Denny Ordway"	"Dean Katz"
## [82]	"Nathan Gelder"	"Mike Vittorini"	"Jack Garza"
## [85]	"Bart Pistole"	"Victor Preis"	"Saphhira Shifley"
## [88]	"Anna Gayman"	"Luke Foster"	"Stephanie Phelps"
## [91]	"Roy FranzÄsisch"	"Keith Herrera"	"Kimberly Carter"
## [94]	"Caroline Jumper"	"Philip Brown"	"Michael Paige"
## [97]	"Natalie Fritzler"	"Janet Molinari"	"Shirley Daniels"
## [100]	"Ken Heidel"	"Ross Baird"	"Dave Brooks"
## [103]	"Mike Kennedy"	"Philisse Overcash"	"Brenda Bowman"
## [106]	"Troy Blackwell"	"Raymond Buch"	"Ed Braxton"
## [109]	"Sanjit Chand"	"Joseph Holt"	"Tanja Norvell"
## [112]	"Joni Sundaesam"	"Craig Carreira"	"Maya Herman"
## [115]	"Jeremy Pistek"	"Kristen Hastings"	"Jeremy Ellison"
## [118]	"John Grady"	"Xylona Preis"	"Erin Mull"
## [121]	"Michelle Tran"	"Sue Ann Reed"	"Carl Weiss"
## [124]	"Astrea Jones"	"Sonia Sunley"	"Rose O'Brian"
## [127]	"Maribeth Dona"	"Greg Tran"	"Ashley Jarboe"
## [130]	"Olvera Toch"	"Liz Pelletier"	"Cynthia Arntzen"
## [133]	"Jeremy Farry"	"Frank Preis"	"Ellis Ballard"
## [136]	"Jennifer Ferguson"	"Sarah Foster"	"Trudy Glocke"
## [139]	"Carlos Soltero"	"Christina Anderson"	"Sylvia Foulston"
## [142]	"Meg O'Connel"	"Annie Thurman"	"Logan Currie"
## [145]	"Fred McMath"	"Brian Dahlen"	"Denny Joy"
## [148]	"Max Engle"	"Rick Bensley"	"Justin Deggeller"
## [151]	"John Lee"	"Sean Christensen"	"Tamara Willingham"
## [154]	"Chuck Clark"	"Barry Blumstein"	"Anthony Rawles"
## [157]	"Steven Roelle"	"Craig Reiter"	"Eugene Hildebrand"
## [160]	"Cassandra Brandow"	"Carol Darley"	"Doug Jacobs"
## [163]	"Grant Thornton"	"Michael Chen"	"Ralph Arnett"
## [166]	"Naresj Patel"	"Alan Barnes"	"Jesus Ocampo"
## [169]	"Dorris liebe"	"Jay Kimmel"	"Brad Norvell"
## [172]	"David Philippe"	"Tracy Hopkins"	"Arthur Prichep"
## [175]	"Roland Schwarz"	"Seth Vernon"	"Steve Chapman"
## [178]	"Jay Fein"	"Emily Grady"	"Darrin Sayre"
## [181]	"Gary Zandusky"	"Phillina Ober"	"Sung Shariari"
## [184]	"Helen Wasserman"	"Mike Caudle"	"Gary McGarr"
## [187]	"Russell Applegate"	"Allen Arnold"	"Zuschuss Carroll"
## [190]	"Cyma Kinney"	"Dave Poirier"	"Linda Southworth"
## [193]	"Guy Phonely"	"Paul Knutson"	"Sally Matthias"
## [196]	"Guy Armstrong"	"Anthony Garverick"	"Dave Kipp"
## [199]	"Peter McVee"	"Lauren Leatherbury"	"Jill Stevenson"

## [202]	"Ed Ludwig"	"Pamela Coakley"	"Hunter Lopez"
## [205]	"Maribeth Schnelling"	"George Bell"	"Justin Ritter"
## [208]	"Bill Eplett"	"Sample Company A"	"Brian Moss"
## [211]	"Katherine Hughes"	"Greg Hansen"	"Scott Williamson"
## [214]	"Joseph Airdo"	"Daniel Lacy"	"Lindsay Williams"
## [217]	"Thomas Brumley"	"Bryan Spruell"	"Bart Watters"
## [220]	"Robert Waldorf"	"Tracy Zic"	"Ann Steele"
## [223]	"Toby Swindell"	"Sara Luxemburg"	"Troy Staebel"
## [226]	"Roland Fjeld"	"Mitch Willingham"	"Rob Dowd"
## [229]	"Frank Merwin"	"Rick Hansen"	"Philip Fox"
## [232]	"John Stevenson"	"Kalyca Meade"	"Nicole Fjeld"
## [235]	"Alyssa Tate"	"Cathy Armstrong"	"Rob Beeghly"
## [238]	"Harold Ryan"	"Bradley Talbott"	"Nat Gilpin"
## [241]	"Larry Hughes"	"Steven Ward"	"Stefania Perrino"
## [244]	"Ben Ferrer"	"Becky Martin"	"Kean Thornton"
## [247]	"Brooke Gillingham"	"Patrick Gardner"	"Christine Sundaesam"
## [250]	"Chris Cortes"	"Phillip Flathmann"	"Dan Campbell"
## [253]	"Sibella Parks"	"Maurice Satty"	"Bryan Davis"
## [256]	"Karen Carlisle"	"Toby Ritter"	"Gene McClure"
## [259]	"Alan Dominguez"	"Todd Boyes"	"Henry Goldwyn"
## [262]	"Laurel Workman"	"Matt Collins"	"Liz Preis"
## [265]	"Evan Bailliet"	"George Zrebassa"	"Laurel Elliston"
## [268]	"Cathy Prescott"	"Frank Gastineau"	"Lisa DeCherney"
## [271]	"Alejandro Ballentine"	"Michael Nguyen"	"Jim Radford"
## [274]	"Jamie Frazer"	"Chad McGuire"	"Aaron Smayling"
## [277]	"Julie Kriz"	"Paul Prost"	"Damala Kotsonis"
## [280]	"Liz Thompson"	"Shui Tom"	"Bill Donatelli"
## [283]	"Michael Moore"	"Pauline Webber"	"Shaun Chance"
## [286]	"Nicole Hansen"	"Mick Brown"	"Linda Cazamias"
## [289]	"Keith Dawkins"	"Maria Bertelson"	"Odella Nelson"
## [292]	"Katherine Nockton"	"Neoma Murray"	"Nora Pelletier"
## [295]	"Catherine Glotzbach"	"Cindy Stewart"	"Katherine Ducich"
## [298]	"Don Weiss"	"Christine Phan"	"Maxwell Schwartz"
## [301]	"Corinna Mitchell"	"Corey-Lock"	"Dionis Lloyd"
## [304]	"Todd Sumrall"	"Jane Waco"	"John Dryer"
## [307]	"Neil Knudson"	"Lindsay Shagiari"	"Jessica Myrick"
## [310]	"Arianne Irving"	"Tom Boeckenhauer"	"Victoria Wilson"
## [313]	"Paul Lucas"	"Gary Hwang"	"Benjamin Venier"
## [316]	"Mike Pelletier"	"Jim Mitchum"	"Carl Ludwig"
## [319]	"Chloris Kastensmidt"	"Deborah Brumfield"	"Emily Phan"
## [322]	"Maria Etezadi"	"Ivan Gibson"	"Bobby Odegard"
## [325]	"Aimee Bixby"	"Julia West"	"Edward Becker"
## [328]	"Sheri Gordon"	"Charlotte Melton"	"Anthony Witt"
## [331]	"Dianna Vittorini"	"Adam Hart"	"Mick Crebagga"
## [334]	"Tonja Turnell"	"Anne Pryor"	"Dan Lawera"
## [337]	"Vivek Grady"	"Muhammed Yedwab"	"Rick Duston"
## [340]	"Dennis Pardue"	"Maribeth Yedwab"	"Jas O'Carroll"
## [343]	"Marina Lichtenstein"	"Yana Sorensen"	"Parhena Norris"
## [346]	"Jenna Caffey"	"Christina DeMoss"	"Thea Hendricks"
## [349]	"Michelle Moray"	"Quincy Jones"	"Neola Schneider"
## [352]	"Robert Barroso"	"Shaun Weien"	"Eric Barreto"
## [355]	"Sean Braxton"	"Pamela Stobb"	"Claudia Bergmann"
## [358]	"Herbert Flentye"	"Henia Zydlo"	"Steven Cartwright"
## [361]	"Natalie DeCherney"	"Cynthia Delaney"	"Jamie Kunitz"

## [364]	"Debra Catini"	"Bradley Nguyen"	"Filia McAdams"
## [367]	"Emily Ducich"	"Bryan Mills"	"Dianna Arnett"
## [370]	"Joni Wasserman"	"Raymond Messe"	"Max Ludwig"
## [373]	"Darren Powers"	"Jill Fjeld"	"Adrian Shami"
## [376]	"Phillip Breyer"	"Frank Olsen"	"Brian Thompson"
## [379]	"Janet Lee"	"Cathy Hwang"	"Neil Ducich"
## [382]	"Bruce Stewart"	"Barbara Fisher"	"Katharine Harms"
## [385]	"Jennifer Braxton"	"Noah Childs"	"Charles Crestani"
## [388]	"Michelle Ellison"	"Benjamin Patterson"	"Mark Packer"
## [391]	"John Castell"	"James Galang"	"Erica Hernandez"
## [394]	"Michael Dominguez"	"Roy Phan"	"Rick Wilson"
## [397]	"Duane Benoit"	"Frank Hawley"	"Erica Hackney"
## [400]	"Helen Abelman"	"Edward Hooks"	"Mary Zewe"
## [403]	"Scot Coram"	"Greg Matthias"	"Valerie Mitchum"
## [406]	"Joe Elijah"	"Ann Chong"	"Dean Braden"
## [409]	"Joy Daniels"	"Kean Takahito"	"Christy Brittain"
## [412]	"Adam Shillingsburg"	"Clytie Kelty"	"Theone Pippenger"
## [415]	"Greg Guthrie"	"Patrick Bzostek"	"Arthur Wiediger"
## [418]	"Roger Barcio"	"Maris LaWare"	"Dorothy Badders"
## [421]	"Matt Hagelstein"	"Dennis Kane"	"Bobby Trafton"
## [424]	"Alex Avila"	"Denny Blanton"	"John Huston"
## [427]	"Carol Triggs"	"Rick Huthwaite"	"Sanjit Jacobs"
## [430]	"Toby Gnade"	"Barry Franz"	"Justin MacKendrick"
## [433]	"Sean Miller"	"Tom Prescott"	"Liz MacKendrick"
## [436]	"Jim Karlsson"	"Patrick Jones"	"Ben Wallace"
## [439]	"Ricardo Sperren"	"Theresa Swint"	"Michael Stewart"
## [442]	"Susan Vittorini"	"Becky Castell"	"Elizabeth Moffitt"
## [445]	"Bradley Drucker"	"Brendan Murry"	"Kristina Nunn"
## [448]	"Kelly Williams"	"Barry FranzÄsisch"	"Katherine Murray"
## [451]	"Brad Eason"	"Christina VanderZanden"	"Speros Goranitis"
## [454]	"Tamara Manning"	"Eleni McCrary"	"Michelle Lonsdale"
## [457]	"Matt Collister"	"Chuck Sachs"	"Carlos Meador"
## [460]	"Greg Maxwell"	"Alan Hwang"	"Tim Brockman"
## [463]	"John Murray"	"Sonia Cooley"	"Bill Tyler"
## [466]	"Luke Schmidt"	"Ralph Ritter"	"Richard Bierner"
## [469]	"Daniel Byrd"	"Nora Preis"	"Peter BÄhler"
## [472]	"Dorothy Wardle"	"Thomas Thornton"	"Lori Olson"
## [475]	"Ken Dana"	"Nicole Brennan"	"Sanjit Engle"
## [478]	"Edward Nazzal"	"Deirdre Greer"	"Craig Leslie"
## [481]	"Hunter Glantz"	"Nancy Lomonaco"	"Rick Reed"
## [484]	"Toby Carlisle"	"Stewart Visinsky"	"Mick Hernandez"
## [487]	"Yoseph Carroll"	"Charles McCrossin"	"Anthony Johnson"
## [490]	"Bobby Elias"	"Resi PÄlking"	"Julie Prescott"
## [493]	"Dean percer"	"Michael Oakman"	"Jennifer Jackson"
## [496]	"Rob Williams"	"Becky Pak"	"Lisa Hazard"
## [499]	"Eugene Moren"	"Michelle Arnett"	"Andy Yotov"
## [502]	"Michael Grace"	"Joy Bell-"	"Amy Cox"
## [505]	"Zuschuss Donatelli"	"Karen Ferguson"	"Tony Chapman"
## [508]	"Sean Wendt"	"Hilary Holden"	"Jim Epp"
## [511]	"Eugene Barchas"	"Chad Cunningham"	"Craig Yedwab"
## [514]	"Cari Sayre"	"Ross DeVincentis"	"Suzanne McNair"
## [517]	"Tamara Chand"	"Justin Ellison"	"Daniel Raglin"
## [520]	"Jack Lebron"	"Fred Hopkins"	"Roy Collins"
## [523]	"Sarah Jordon"	"Bruce Geld"	"Matt Abelman"

## [526]	"Ruben Ausman"	"Laurel Beltran"	"Candace McMahon"
## [529]	"Randy Ferguson"	"Evan Henry"	"Kean Nguyen"
## [532]	"Penelope Sewall"	"Art Foster"	"Jason Fortune-"
## [535]	"Karen Daniels"	"Patrick Ryan"	"Christine Abelman"
## [538]	"Mark Cousins"	"Liz Carlisle"	"Natalie Webber"
## [541]	"Toby Braunhardt"	"Cindy Schnelling"	"Matthew Grinstein"
## [544]	"Dave Hallsten"	"Deanra Eno"	"Sam Zeldin"
## [547]	"Ryan Akin"	"Kelly Andreada"	"Jack O'Briant"
## [550]	"Heather Jas"	"Tony Sayre"	"Robert Dilbeck"
## [553]	"Jim Kriz"	"Aleksandra Gannaway"	"Randy Bradley"
## [556]	"Corey Catlett"	"Darren Budd"	"Vicky Freymann"
## [559]	"Marc Crier"	"Jennifer Patt"	"Jasper Cacioppo"
## [562]	"Maureen Fritzler"	"Brendan Sweed"	"William Brown"
## [565]	"Michelle Huthwaite"	"Dario Medina"	"Joel Jenkins"
## [568]	"Anna HÃ¼berlin"	"Harry Greene"	"Craig Carroll"
## [571]	"Lycoris Saunders"	"Kunst Miller"	"Katrina Edelman"
## [574]	"Duane Huffman"	"Gary Hansen"	"Darrin Martin"
## [577]	"Nathan Cano"	"Peter Fuller"	"Stuart Calhoun"
## [580]	"Valerie Takahito"	"Adrian Barton"	"Dana Kaydos"
## [583]	"Khloe Miller"	"Susan Pistek"	"George Ashbrook"
## [586]	"Darren Koutras"	"Lena Creighton"	"Lynn Smith"
## [589]	"Matthew Clasen"	"Lena Hernandez"	"Thomas Boland"
## [592]	"Dorothy Dickinson"	"Ann Blume"	"Ted Trevino"
## [595]	"Paul Van Hugh"	"Eileen Kiefer"	"Joe Kamberova"
## [598]	"Cyra Reiten"	"Susan Gilcrest"	"Neil FranzÃ¶sisch"
## [601]	"Maria Zettner"	"Shirley Jackson"	"Anne McFarland"
## [604]	"Stefanie Holloman"	"Tiffany House"	"Magdelene Morse"
## [607]	"Mark Van Huff"	"Alex Grayson"	"Nat Carroll"
## [610]	"Beth Thompson"	"Paul Stevenson"	"Trudy Brown"
## [613]	"Mitch Webber"	"Steve Nguyen"	"Mark Hamilton"
## [616]	"Erin Creighton"	"David Wiener"	"Meg Tillman"
## [619]	"Anna Andreadi"	"Noel Staavos"	"Giulietta Weimer"
## [622]	"Georgia Rosenberg"	"Vivek Sundaresam"	"Tom Stivers"
## [625]	"Michael Granlund"	"Alejandro Grove"	"Richard Eichhorn"
## [628]	"Nick Radford"	"Christine Kargatis"	"Maureen Gnade"
## [631]	"Liz Willingham"	"Mary O'Rourke"	"Karen Seio"
## [634]	"Ivan Liston"	"Denise Leinenbach"	"Allen Goldenen"
## [637]	"Lisa Ryan"	"Shahid Shariari"	"Vivek Gonzalez"
## [640]	"Hallie Redmond"	"Scott Cohen"	"James Lanier"
## [643]	"Cari Schnelling"	"Andrew Gjertsen"	"Chris Selesnick"
## [646]	"Harry Marie"	"Pauline Johnson"	"Bill Overfelt"
## [649]	"Tony Molinari"	"Stephanie Ulpright"	"Brian DeCherney"
## [652]	"Aaron Hawkins"	"Andy Reiter"	"Carlos Daly"
## [655]	"Robert Marley"	"Don Miller"	"Tom Ashbrook"
## [658]	"Matt Connell"	"Bruce Degenhardt"	"Clay Rozendal"
## [661]	"Adrian Hane"	"Jason Gross"	"Amy Hunt"
## [664]	"Max Jones"	"Doug O'Connell"	"Brian Stugart"
## [667]	"Chris McAfee"	"Ruben Dartt"	"Barry Pond"
## [670]	"Jill Matthias"	"Roger Demir"	"Joy Smith"
## [673]	"Erica Smith"	"Scot Wooten"	"Brian Derr"
## [676]	"Angele Hood"	"Evan Minnotte"	"Anthony O'Donnell"
## [679]	"Anna Chung"	"Mark Haberlin"	"Giulietta Baptist"
## [682]	"Alan Shonely"	"Katrina Bavinger"	"Annie Zypern"
## [685]	"Sarah Bern"	"Tracy Poddar"	"Nona Balk"

```
## [688] "Beth Fritzler"      "Frank Carlisle"      "Dennis Bolton"
## [691] "Pauline Chand"      "Anthony Jacobs"      "Shirley Schmidt"
## [694] "Lela Donovan"       "David Kendrick"      "Carol Adams"
## [697] "Ed Jacobs"          "MaryBeth Skach"      "Juliana Krohn"
## [700] "Victoria Pisteka"   "Paul MacIntyre"      "Ted Butterfield"
## [703] "Barry Gonzalez"     "Rob Lucas"           "Charles Sheldon"
## [706] "Nick Crebassa"      "Giulietta Dortch"    "Russell D'Ascenzo"
## [709] "Thea Hudgings"      "Henry MacAllister"   "Anemone Ratner"
## [712] "David Flashing"     "Eva Jacobs"          "Monica Federle"
## [715] "Marc Harrigan"      "Lena Radford"        "Mike Gockenbach"
## [718] "Jason Klamczynski"  "Bill Stewart"        "Craig Molinari"
## [721] "Denise Monton"      "Shahid Collister"    "Eudokia Martin"
## [724] "Stuart Van"         "Alex Russell"        "Steve Carroll"
## [727] "Elpida Rittenbach"  "Luke Weiss"          "Lena Cacioppo"
## [730] "Trudy Schmidt"      "Lindsay Castell"     "Sharelle Roach"
## [733] "Sarah Brown"        "Harold Engle"        "Christopher Conant"
## [736] "Art Ferguson"       "Aaron Bergman"       "Bill Shonely"
## [739] "Skye Norling"       "Mathew Reese"        "Nora Paige"
## [742] "Julia Barnett"      "Brendan Dodson"      "Ritsa Hightower"
## [745] "Carl Jackson"       "Harold Dahlen"       "Beth Paige"
## [748] "Cindy Chapman"      "Fred Wasserman"      "Ionia McGrath"
## [751] "Roy Skaria"         "Tracy Collins"       "Alice McCarthy"
## [754] "Erin Smith"         "Sung Chung"          "Eric Murdock"
## [757] "Christopher Martinez" "Vivian Mathis"      "Benjamin Farhat"
## [760] "Rachel Payne"       "Sam Craven"          "Susan MacKendrick"
## [763] "Brad Thomas"       "Ryan Crowe"

find_bill_df <- stringr::str_split_fixed(string=sales_pipe$Customer.Name,pattern = " ", n=2)

length(find_bill_df [find_bill_df == "Bill"])

## [1] 37
```

Exercise 6

How many mentions of the word 'table' are there in the Product.Name column? **Note you can do this in one line of code**

```
# Your code here

sum(stringr::str_count(sales_pipe$Product.Name, "table"))

## [1] 240
```

Exercise 7

Create a table of counts for each state in the sales data. The counts table should be ordered alphabetically from A to Z.

```
# Your code here

table(sales_pipe$State)

##
## Alabama Arizona Arkansas
```

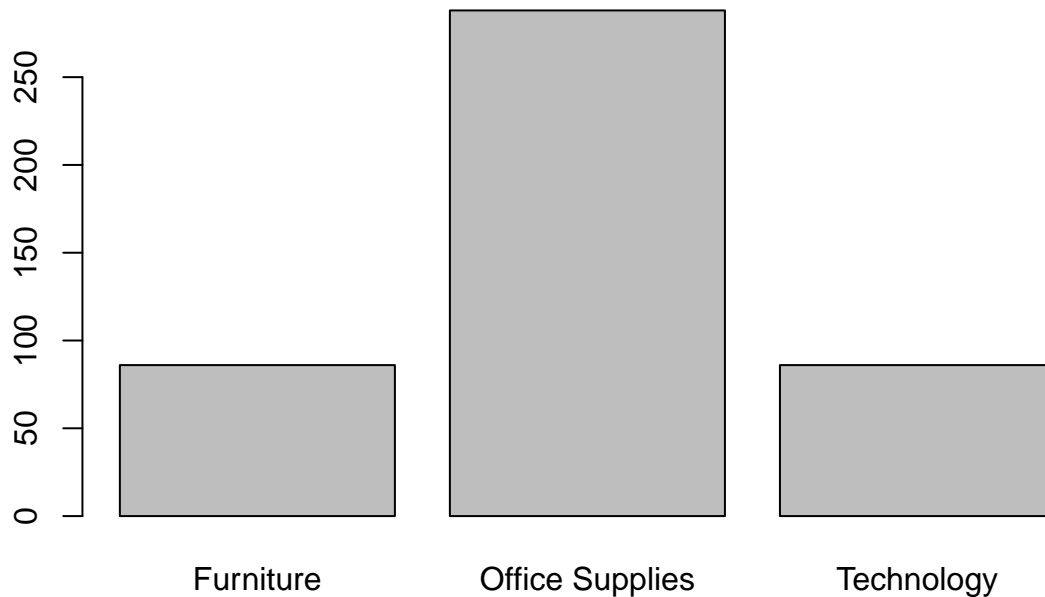

##	28	119	22
##	California	Colorado	Connecticut
##	993	90	50
##	Delaware	District of Columbia	Florida
##	47	1	186
##	Georgia	Idaho	Illinois
##	79	9	286
##	Indiana	Iowa	Kansas
##	74	11	16
##	Kentucky	Louisiana	Maine
##	64	18	4
##	Maryland	Massachusetts	Michigan
##	63	71	142
##	Minnesota	Mississippi	Missouri
##	41	27	37
##	Montana	Nebraska	Nevada
##	2	26	24
##	New Hampshire	New Jersey	New Mexico
##	9	58	11
##	New York	North Carolina	North Dakota
##	555	117	7
##	Ohio	Oklahoma	Oregon
##	211	38	56
##	Pennsylvania	Rhode Island	South Carolina
##	312	25	28
##	South Dakota	Tennessee	Texas
##	9	88	460
##	Utah	Vermont	Virginia
##	27	10	80
##	Washington	West Virginia	Wisconsin
##	254	4	38
##	Wyoming		
##	1		

Exercise 8

Create an alphabetically ordered barplot for each sales Category in the State of Texas.

Your code here

```
texas_df = sales_pipe[sales_pipe$State=='Texas',c("Category")]
barplot(table(texas_df))
```



Exercise 9

Find the average profit by region. **Note:** You will need to use the `aggregate()` function to do this. To understand how the function works type `?aggregate` in the console.

Your code here

```
aggregate(sales_pipe$Profit, list(sales_pipe$Region), FUN=mean)
```

```
##   Group.1      x
## 1 Central 20.46822
## 2   East 29.91937
## 3  South 11.27720
## 4   West 32.77000
```

Exercise 10

Find the average profit by order year. **Note:** You will need to use the `aggregate()` function to do this. To understand how the function works type `?aggregate` in the console.

Your code here

```
order.date.split <- stringr::str_split_fixed(sales_pipe$Order.ID, pattern="-", n=3)
```

```
order.date.yr <- order.date.split[,2]
```

```
aggregate(sales_pipe$Profit, list(order.date.yr), FUN=mean)
```

```
##   Group.1      x  
## 1    2014 32.24582  
## 2    2015 21.58676  
## 3    2016 30.10960  
## 4    2017 21.31825
```